

REPASO ESTADÍSTICAS, INVESTIGACIÓN Y MEDICIÓN

Mayo 2011

Preparado por:

Armando Torres Pesaresi, M.S.
Estudiante Doctoral de Psicología Clínica (Psy.D.)

Editado por:

Javier I. Toro Torres, Ph.D.
Psicólogo Clínico

REPASO ESTADÍSTICAS, INVESTIGACIÓN, CONSTRUCCIÓN DE PRUEBAS Y MEDICIÓN

Curva Normal: Las puntuaciones de la inmensa mayoría de las personas se amontonan en la gama o amplitud intermedia. Un pequeño número se agrupa hacia los extremos elevado y bajos. Los dos lados son simétricos. Y la media, la moda y la mediana coninciden entre sí.

Curva Sesgada o Asimétrica: Cuando la media, la moda y la mediana no son iguales.

Efecto Halo: Características dominante en un individuo que puedan afectar a la impresión percibida de otras que posee. Ej. Todos los judíos son materialistas; un niño está por primera vez con su maestra y saca F en el primer examen, esto hace que la maestra piense que el niño no es aplicado en sus estudios.

Taxonomía de Bloom:

- **Conocimiento:** Retención del material aprendido.
- **Comprensión:** Entender el significado del material.
- **Aplicación:** Usar la información en situaciones concretas
- **Síntesis:** Integrar las partes en un todo.
- **Evaluación:** Juzgar el valor de algo para un propósito.

Validez: El grado en que una prueba y/o instrumento mide lo que se supone que mida.

- **Validez de Contenido:** Se refiere al contenido y formato del instrumento. El contenido y el formato deben ser consistentes con la definición de la variable y la muestra de sujetos a ser evaluados. Se tiene que determinar si el contenido es una muestra adecuada del dominio del contenido el cual está supuesto a representar. También tiene que ver con el formato del instrumento, como por ejemplo: la claridad en la copia, tamaño de la letra, claridad en las direcciones, adecuación del lugar de administración, si el formato es apropiado entre otros. Si el formato se presenta de forma inadecuada, no se pueden obtener resultados válidos.
- **Validez de Criterio:** Se refiere a la relación entre las puntuaciones obtenidas utilizando el instrumento y puntuaciones obtenidas utilizando uno o más instrumentos o medidas. Un criterio es un segundo instrumento el cual se presume mide la misma variable. ¿Cuán fuerte es la relación?

Existen dos tipos de Validez de Criterio:

Validez Predictiva: Se determina un período de tiempo entre administraciones del instrumento y se obtiene la puntuación de criterio.

Validez Concurrente: Cuando la data de un instrumento se obtiene más o menos al mismo tiempo, y se comparan los resultados obtenidos.

- **Validez de Constructo:** Se refiere a la naturaleza del constructo psicológico o característica a ser medida por el instrumento. Existen tres pasos en la obtención

de este tipo de validez: Variable a ser medida se encuentra claramente definida, Las hipótesis, basadas en la teoría, se forman en cómo la persona que posee “mucho” versus la que posee “poco” de la variable se va a comparar en una situación en particular. Las hipótesis es probada lógica y empíricamente.

Escalas de Medición:

- **Nominal:** En esta escala la variable se desglosa en categorías que identifican particularidades. (los números no refleja un orden). Ej. Religión, género, idioma, etc.
- **Ordinal:** Los números en esta escala reflejan el orden de los sujetos, no la magnitud de las diferencias. Las diferencias entre el uno, el dos y el tres no son iguales. Cuando pones algo por predilección de gustos por ejemplo pero puede ser que entre la espinaca, el helado y el cheese cake; el más que te guste sea el cheese cake, después el helado y por último la espinaca.
- **Intervalo:** Estas escalas proporcionan números que reflejan tanto orden como las diferencias entre los sujetos, pero estas escalas carecen de un cero absoluto. Ej. Inteligencia, Temperatura.
- **Razón:** Iguales a las escalas de intervalo, pero tienen cero absoluto. Ej. Distancia.

Prueba de Aptitud: Determinan si el individuo posee las facultades o el talento latente para aprender un determinado oficio.

Escala de Actitudes: Estado de disposición psicológica adquirida y organizada a través de la propia experiencia que incita al individuo a reaccionar de una manera característica frente a determinadas personas, objetivos o situaciones.

Un **Prueba Psicológica** es una medida objetiva tipificada de una muestra de conducta. Es una medida objetiva porque es de criterio cuantitativo. Tipificada o Estandarizada porque ha seguido una metodología rigurosa, según el manual. No podemos alterar una prueba o subprueba esto es una falta de ética. Una muestra de conducta porque cuando hagamos una evaluación vamos a tener información sobre un aspecto del comportamiento de ese individuo.

Examen Estandarizado: Es una prueba aplicada a una muestra representativa de personas que la responden con el propósito de establecer normas (proceso de estandarización). Cuando una prueba está *estandarizada* tiene procedimientos definidos en forma clara para su administración y calificación, incluyendo datos normativos.

Error Típico de la Medida: Se refiere al grado de confianza que se deposita en un puntaje obtenido de la administración de un examen a un sujeto. Es otra manera de expresar la confiabilidad del examen. Mientras más bajo más confiable.

Medición vs. Evaluación: La diferencia consiste en que la *medición* es objetiva. Además la valoración que se le atribuye a las respuestas es producto del consenso de los investigadores que desarrollaron la prueba. *Evaluar* sigue unos criterios subjetivos, que emite juicio sobre algo, alguien, objeto o situación, basándose este juicio en algunos criterios preestablecidos.

Tipos de Examen:

- **Grupales:** Son pruebas que se les pueden administrar a un grupo de personas a la misma vez.
- **Individuales:** Se aplican a una persona a la vez, requieren la atención casi total del examinador al examinado.
- **Instrumentales:** Son aquellas pruebas que requieren el uso de determinados instrumentos y suelen ser de carácter individual.
- **Pruebas de Papel y Lápiz:** Es donde el individuo contesta por escrito a una determinada pregunta, y suelen ser grupales:
- **Pruebas de Aptitud:** Determinan si el individuo posee las facultades o el talento latente para aprender un determinado oficio.
- **Perfección:** Estas pruebas miden lo acertadamente que el individuo puede realizar el trabajo, o lo que conoce de él en ese momento.
- **Personalidad o Intereses:** El propósito es valorar las características de la personalidad, o los intereses y gustos del individuo.

Relación entre la psicología y la medición: En la psicología científica se usa la medición como medio para efectuar los análisis y las interpretaciones de la realidad que se estudia. También se usa la evaluación subjetiva para esto.

Inventarios de Personalidad:

- **MMPI**
- **NEO-PI(R):** 5 grandes factores de personalidad
- **16 PF:** Raymond Catell. 16 factores de la personalidad
- **Escala F de California:** Utilizada para medir autoritarismo. Constructo popularizado por Theodore W. Adorno.
- **Técnica de Selección Q:** Técnica para medir constructos de personalidad.

Cuatro Pilares de la Evaluación:

- Utilizar pruebas normalizadas
- Entrevista-Histórica
- La Observación
- La Evaluación Informal

Examen de Criterio: Exámenes que tienen que ser administrados e interpretados siguiendo unas normas o estándares para quienes los utilizan.

Normalización: Proceso riguroso donde se establecen los procedimientos específicos (manuales, administración, selección de la muestra representativa) que nos lleva a desarrollar normas de la prueba. Nunca podemos alterar los procedimientos establecidos en el manual de una prueba. Estos procedimientos han establecido la confiabilidad y validez del instrumento.

Normas: Nos permiten establecer la posición relativa de una persona en comparación a la ejecución típica de la muestra de estandarización. Vamos a establecer las normas con un grupo primero y luego establecemos la posición de una persona que originalmente tomó esa prueba.

Confiabilidad: Se refiere a la precisión (consistencia y estabilidad) de la medida realizada por una prueba. Que mida lo mismo a través del tiempo. *Tipos de Confiabilidad:*

- **Temporal o Prueba/Posprueba:** Grado en el cual las puntuaciones de una prueba son consistentes de una administración de la prueba a otra. Administrar la misma prueba dos veces al mismo grupo, después de transcurrido un tiempo. A través del tiempo mantiene consistencia o estabilidad.
- **Métodos de Formas Equivalentes o Alternas:** Cuando dos o más versiones o pruebas diferentes pero equivalentes se administran al mismo grupo de sujetos, durante un mismo periodo.
- **Consistencia Interna:** Se requiere una aplicación del instrumento. Se dividen los reactivos del instrumento en dos partes iguales. Estas partes son las que se comparan. Se espera que ambas mitades sean similares en contenido y nivel de dificultad. Uno de los métodos más usados es el Kuder-Richardson que determina la consistencia interna.
- **Procedimiento de Split Half o Mitades:** Este procedimiento envuelve el corregir dos mitades (usualmente ítems impares versus pares) del instrumento para cada persona, y luego calculando el coeficiente de correlación para los dos sets de puntuaciones. El coeficiente indica el grado en que las dos mitades de la prueba proveen los mismos resultados y por ende, describen la consistencia interna de la prueba.

La prueba **Matrices Progresivas de Raven** es una prueba no verbal donde se presentan estímulos figurativos, puede ser tanto a nivel individual como grupal. El tiempo de administración es de aproximadamente 15 a 30 minutos. Se administra individual con propósitos clínicos o de selección de personal, se administra grupal para fines de investigación.

Batería de Pruebas para Niños Kaufman: Reduce el contenido verbal. Los autores tratan de disminuir todo contenido que se relaciona con los aspectos culturales. Se administra a niños entre las edades de 2 años y 6 meses hasta 12 años y 5 meses. Toma 45 minutos aproximado con los niños preescolares, y 75 minutos aproximado con los niños de edad escolar.

Dibujo de la Figura Humana: La prueba es corta de tipo no verbal. En esta prueba se mide tanto inteligencia no verbal (madurez intelectual) como indicadores emocionales.

Bayley de Desarrollo Infantil: Evalúa el desarrollo infantil en niños de 2 años a 2 años y medio. Ofrece dos puntuaciones estándar: Un índice de Desarrollo Mental, donde obtienes la Edad Mental y un índice de Desarrollo Psicomotor (Escala Motora).

Escala de Inteligencia Wechsler para Niños-R, P.R.: Contiene 12 sub-pruebas, seis subpruebas verbales y seis subpruebas de ejecución. Las subpruebas verbales son: Información, Semejanza, Aritmética, Vocabulario, Comprensión y Retención de Dígitos. Las subpruebas de ejecución son: Figuras Incompletas, Arreglo de Dibujos, Diseños con Bloques, Composición de Objetos, Claves y Laberintos. *Retención de Dígitos* y *Laberintos* no se usaron para establecer las tablas de C.I. Estas subpruebas son alternas, si alguna prueba no se puede administrar o se invalida la sustituyes. Se deben administrar las 12 escalas para propósitos clínicos. La **estandarización** de la prueba fue con 2,200 participantes de la población escolar. Se administra a niños 6 años a 16 años 11 meses, el lenguaje vernáculo fue el español y no diagnóstica condiciones físicas ni mentales. Un cociente intelectual de 100 es promedio. La puntuación a escala de 10 es promedios con una desviación estándar de 3. Duración con las 10 pruebas de 50 a 75 minutos. Se debe administrar en la secuencia en que se presentan las sub-pruebas. Se compara al niño con él mismo.

Escala de Inteligencia Wechsler para Adultos: Se normaliza para los adultos puertorriqueños en el 1968 y de esta manera proveer una medida de la inteligencia, válida para la población adulta puertorriqueña. Contó con una muestra de 1,127 y con el censo del 1960.

Estanford Binet: No está normalizada para Puerto Rico pero tiene validez buena. Sale para 1905, el propósito de esta prueba fue medir las capacidades intelectuales de los niños escolares y es la primera prueba en psicología con la intención de esta era clasificar a los niños para verlos de forma individual. Es de mucho contenido verbal. Se utiliza desde 2 años y medio hasta adulto. Es por niveles.

Bender: Evalúa los problemas de percepción e integración motora, percepción visual y destrezas motoras. Son importante para evaluar los problemas de aprendizaje y deficiencias neurológicas. Baja ejecución se asocia con retraso en la maduración, poca estimulación del ambiente y dificultades neurológicas.

Vineland: Evalúa los aspectos sociales y personales de la conducta del evaluado. La escala aplica a personas con o sin impedimento desde cero meses hasta los 18 años con 11 meses o un bajo funcionamiento en la etapa adulta. La escala define el comportamiento adaptativo como la ejecución de las actividades diarias requeridas para suficiencia personal y social.

MMPI: Prueba de personalidad que se desarrolla con un fin diagnóstico. Fue diseñada para la administración a adolescentes de 14 años o más. Tiene diez escalas clínicas: Hipocondriasis, Depresión, Histeria, Desviación Psicopática, Masculinidad-Feminidad, Paranoia, Psicastenia, Esquizofrenia, Hipomanía e Introversión Social.

Prueba de Personalidad o Intereses: El propósito es valorar las características de la personalidad, o los intereses y gustos del individuo.

Pruebas de Perfección: Estas pruebas miden lo acertadamente que le individuo puede realizar el trabajo, o lo que conoce de él en ese momento.

En las encuestas se debe evitar hacer preguntas dirigidas

Un estudio que compara grupos de edades 6,16, 36 son estudios transversales

Estadísticas: Conjunto de procedimientos que se utilizan para reunir, medir, clasificar, codificar, computar y resumir información numérica.

Estadísticas Descriptivas: Nos ayuda a describir observaciones. Cuantifican las variables bajo investigación y permite describirlas y resumirlas.

Estadísticas Inferencial: Con esta se pueden hacer inferencias sobre los resultados que se obtienen. Se utiliza para determinar si se puede rechazar o no las hipótesis de investigación; poder inferir si lo que se observa en una muestra puede generalizarse a la población.

Variables: Cualquier característica bajo estudio que toma diferentes valores. Son cosas que se observan. Es una propiedad que adquiere distintos valores. Ej. La luz, el calor, la depresión con la intensidad de la luz, el efecto en la piel, productividad o ejecución académica.

Pueden ser de dos tipos:

Cualitativas: No se pueden medir numéricamente (Ej. Nacionalidad, color de piel, sexo).

Cuantitativas: Tienen valor numérico (edad, precio de un producto, ingresos anuales).

Estas se pueden clasificar en:

- **Variable Continuas:** Que pueden tomar cualquier valor real dentro de un intervalo. Ej. La velocidad de un carro puede ser 80.3 MPH., una persona que mide 72 ½ pulgadas.... etc.
- **Variable Discreta o Categóricas:** Sólo pueden tomar valores enteros (1, 2, 8, -4, etc). Ej. Número de hermanos, genero, estado civil. Este valor no podrá ser 1 ½ persona por ejemplo.

Variables Independiente: Es la que el investigador manipula deliberadamente. La Variable Independiente o Predictora es aquella en la cual los grupos del estudio interfieren, ya sea porque han sido expuesto a diferentes tratamientos o por características inherentes de los grupos. Es la manipulada por el experimentador a fin de observar sus efectos sobre alguna otra variable previamente seleccionada.

Variable Dependiente: Esta es la preseleccionada, donde habrá un efectos por exponerse a la variable independiente. Aspecto de la conducta en el que se esperan encontrar los efectos producidos por los cambios introducidos por el investigador en la variable independiente.

Variable Contaminadora o Extraña: Son fuentes de variación ajenas a la variable independiente, de efectos sistemáticos y no aleatorios sobre la variable dependiente. Son fenómenos complejos que afectan los estudios. Ej. Grado de entusiasmo de parte de los participantes.

Hipótesis: Es la posible contestación a la pregunta de investigación. Se redacta como una negación (nula) o una afirmación (alterna). Requiere de estadísticas inferenciales para ser o no ser rechazada. Se utilizan las siguientes letras para identificar las variables. **Y-Variable Dependiente y X-Variable Independiente.**

Error alfa o de Tipo 1: Se rechaza la Hipótesis Nula

Error beta o de Tipo 2: Se acepta la Hipótesis Nula

Estudio Piloto o Ensayo de la Prueba: Es el ensayo de la prueba con personas similares en aspectos críticos a las personas para las que está diseñada. Por ejemplo si una prueba está diseñada para ayudar en las decisiones respecto a la selección de empleados corporativos con potencial ejecutivo en un determinado nivel, sería apropiado ensayarla con empleados corporativos en el nivel al que está dirigida, y sería inapropiado hacerlo con estudiante de introducción a la psicología. El ensayo de la prueba deberá llevarse a cabo bajo condiciones que sean lo más similares posibles a las condiciones bajo las cuales se aplicará la prueba estandarizada.

Medidas de Tendencia Central:

- **Media (-x):** Promedio aritmético de un grupo de puntuaciones. Suma de las puntuaciones dividida entre el número de puntuaciones. Es la medida de tendencia central más común. Se utiliza en la escala de intervalo o de razón
- **Mediana (Md):** Es el punto en una distribución de datos que está localizado de tal manera que sobre él y bajo él caen exactamente el 50% de los casos. Esto es cuando arreglas los números de menor a mayor. Es el punto en donde la distribución se divide en 2 partes iguales. Se utiliza en la escala ordinal.
- **Moda (Mo):** Valor numérico que se produce con más frecuencia en una distribución de datos. Es el dato que más se repite en una distribución. Se utiliza en la escala nominal.

Universo o Población: Totalidad de individuos, objetos o medidas. Cualquier grupo total, según lo define el investigador.

Muestra: Una parte supuestamente representativa del universo de la cual se obtuvo.

Parámetro: Cualquier característica de la población que sea medible. Ej. Una población puede ser todos los estudiantes de la Universidad de P.R.- una característica medible de esta población puede ser cuántos utilizan un vehículo para llegar a la universidad.

Medidas de Posición No Centrales o Relativa: Permiten conocer otros puntos característicos de la distribución que no son los valores centrales; se dividen en:

- **Cuartiles:** Son 3 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cuatro tramos iguales, en los que cada uno de ellos concentra el 25% de los resultados. Ej. Un dólar dividido en pesetas, son 4 pesetas que representan el 25%.
- **Deciles:** Son 9 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en diez tramos iguales, en los que cada uno de ellos se concentra el 10% de los resultados. Ej. Es dividir un dólar en diez centavos.
- **Percentiles:** Son 99 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cien tramos iguales, en los que cada uno de ellos concentra el 1% de los resultados. Ej. Divides un dólar en centavos. A mayor percentil mejor sale una persona.
- **Rango Percentil:** Compara un resultado del estudiante con una muestra nacional de estudiantes en el mismo grado. El rango percentil indica el porcentaje de estudiantes en la nación que lograron una puntuación igual o más baja que su hijo. Por ejemplo, Ashley con un rango percentil de 38% logró una puntuación mejor que el 38% de los estudiantes que examinaron de ese grado. El resto de los estudiantes, 62% anotaron igual o superaron a su hija.

Muestreo: Se refiere al método usado para seleccionar un número de personas o cosas de una población. La estrategia utilizada influye en la calidad de los datos y en las inferencias que se puedan hacer a partir de ellos.

Muestreo Aleatorio: Todos los miembros de la población tienen la misma oportunidad de ser seleccionados. La idea básica es que los individuos seleccionados sean tan similares como los que no se seleccionaron.

- **Muestreo Aleatorio Simple:** Muestra elegida independientemente de todas las demás, con la misma probabilidad que cualquier otro y cuyos elementos están elegidos independientemente unos de otros y con la misma probabilidad. Es el mejor tipo de muestra y produce muestras representativas.
- **Muestreo Aleatorio Estratificado:** Se subdivide la población en categorías o estratos y se escoge X número de sujetos de cada categoría en la misma proporción que aparecen en la población. Esta aumenta la probabilidad de representatividad, especialmente si la muestra no es muy grande. Ej. En una población de 365 estudiantes donde existen 219 mujeres, que representan el 60% y 146 varones que representan un 40%. Se dividirá en 66 mujeres que conforman un 60% y 44 varones que conforman un 40%.
- **Muestreo Aleatorio por Conglomerado (Cluster Random Sampling):** Se obtiene la muestra al azar utilizando grupos como unidad de muestreo y no sujetos. Pueden ser escuela, encuestas políticas, calles, etc. La ventaja es que se puede usar en las situaciones que son difícil o imposible seleccionar muestras de sujetos. Además, es más fácil de implantar en escuelas, y conlleva menos tiempo. Ej selecciono la escuela al azar y después los grupos al azar, si tengo 100 salones selecciono 10 salones al azar que vendría a ser mi muestra. Pero cada salón tiene 20 estudiantes; yo saco un promedio por salón lo uso como uno y después calculo el promedio de 10 salones.

Muestreo No Aleatorio: Cada miembro de la población no tiene la misma posibilidad de ser seleccionado, en ocasiones no tiene ninguna posibilidad.

- **Muestreo Sistemático:** Se selecciona un número de sujetos de la lista de la población. Su desventaja es que no todos tienen la misma oportunidad porque la población puede estar ordenada sistemáticamente.
- **Muestra por Conveniencia o Disponibilidad:** Muchas veces es difícil o imposible seleccionar muestras aleatorias o sistemáticas. Estas consisten de los grupos de sujetos que están convenientemente disponibles para el estudio. La ventaja es que es conveniente y está disponible; la desventaja es de que hay una tendencia a estar prejuiciada. No se considera representativa de la población y debe evitarse de ser posible.
- **Muestreo a Propósito o Intencional:** Consiste en seleccionar la muestra basado en la experiencia y el conocimiento previo del investigador. Este usa un juicio personal para seleccionar la muestra.

Error de Muestreo: Cuando se aplican estrategias ideales de muestreo y se forman al azar grandes muestra representativas de la población, los estudios sobre las muestras producirán resultados un tanto diferentes de los que se obtienen estudiando toda la

población. Esto se debe a las diversas variaciones en los individuos que se escogieron para la investigación.

Gráfica: Es un diagrama o esquema compuesto de líneas, puntos, barras u otros símbolos que describen los datos que se están midiendo.

- **Histograma:** Es una gráfica con líneas verticales trazadas en los límites verdaderos de cada puntuación de prueba que forma una serie de rectángulos contiguos. Se acostumbra colocar las puntuaciones de la prueba a lo largo del eje horizontal de la gráfica y los números que indican frecuencia de ocurrencia se colocan a lo largo del eje vertical de la gráfica.
- **Polígono de Frecuencia:** Los datos ilustrados en un polígono de frecuencia se expresan por una línea continua que conecta los puntos donde se encuentran las puntuaciones de prueba o intervalos de clase.
- **Ojiva o Curva de Porcentaje Acumulado:** En esta se grafica con unidades de frecuencias acumuladas y no de porcentajes acumulados. Los percentiles y los rangos percentiles se pueden leer directamente a partir de la curva de porcentajes acumulados. Se llama ojiva que implica una forma de S. Es una gráfica para hallar medidas de posición relativa.

Números Relativos: Permiten al investigador establecer relaciones simples, lo que facilita no solo el extraer conocimientos y las referencias numéricas, sino también hacer las comparaciones sin los inconvenientes asociados a los números.

- **Razón:** Es utilizada con frecuencia. Relaciona 2 cantidades. Ej. 8 mujeres por cada 10 hombres.
- **Proporción:** La suma de proporciones en una distribución dada es igual a 1. Relaciona 3 cantidades o más. Ej. $P = A/A+B$, donde P = proporción, A = cantidad que se quiere cambiar a proporción y B = a la otra cantidad.
- **Por Porcentaje:** Proporciones por 100.
- **Cambio Porcentual:** Es una medida que permite determinar en cuanto por ciento ha mermado o aumentado un fenómeno entre 2 puntos de referencia. Es ver el cambio a través del tiempo.
- **Tasas:** Se relaciona a un evento con la población expuesta a dicho evento. Las mismas indican riesgo. Se afectan por: distribución de población por edad, ingreso y educación. Para eliminar 3 puntos decimales o más, las tasa se multiplican por 1000.

Distribución de Frecuencias: Agrupa los datos para facilitar el manejo. Organiza la información en forma lógica para que se le puedan aplicar los métodos estadísticos.

- **Tablas Simples:** Contienen una variable. Presentan en la columna izquierda la característica que se intenta estudiar, acompañada de otra columna a la derecha indicando frecuencia o número de participantes que caen en esa categoría.
- **Tablas Bivariadas:** Permiten observar la posible existencia de relaciones entre características bajo estudio. Se cruzan 2 variables.
- **Tablas Multivariadas:** Combina 3 o más variables. Facilitan mantener constante el efecto de terceras variables.

- a) **Arreglo de Valores:** Se ubican los números en un orden particular. Generalmente en forma descendente.
- b) **Clases no Agrupadas:** Quiere decir el número de términos incluidos o los valores no agrupados.
- c) **Clases Agrupadas:** Cuando se suman más cantidades de intervalos (suma de las $f = \sum f$). Se logra sumar las frecuencias a partir de la clase menor en dirección a la mayor. Cuando se agrupa es conveniente mantener la distribución en las clases: no menos de 10 ni mayor de 20 siempre que sea posible debe proveerse un intervalo igual para todas las clases.
 1. **Límites Aparentes:** Son los puntos donde comienza y terminan las clases. Este número es dado.
 2. **Punto Medio:** Se suman los límites aparentes de los intervalos y se divide entre 2. El valor en la clase agrupada que ha de tomar como representativo de la clase. Es la puntuación media del intervalo.
 3. **Intervalo de Clase:** El número de los términos incluidos en cada clase. Recorrido de cada clase. El número de clases es el número inferior del intervalo menos el número inferior del otro intervalo.

Medidas de Variación o de Dispersión: Mientras menor es la puntuación de variabilidad mayor es el parecido al grupo que estoy midiendo.

- **Recorrido:** Es el índice más simple y se define como la diferencia entre el dato de mayor y el de menor valor numérico. Es un dato para comparar cuán amplio fue el recorrido de una distribución vs. la(s) otra(s).
- **Varianza:** Representa lo que conocemos como la homogeneidad (que se parecen) y la heterogeneidad (que difieren) de los datos. Los datos más homogéneos tendrán una varianza pequeña y los heterogéneos una más grande. La varianza para la muestra es S^2 y para el universo σ^2 .
- **Desviación Estándar:** Es una medida de referencia para saber cuánto difieren los datos y cuanto se asemejan. Es un tipo de agrupación que considera el nivel de separación de los datos. El valor de la desviación estándar se sumará y restará al valor obtenido de la medida aritmética para así comparar la dispersión de los datos a comparar. La DS para la muestra es S y para el universo es σ . : Es la raíz cuadrada de la varianza. Para computar la desviación estandar se utilizan todos los puntajes y ésta se considera una de las mejores medidas de variabilidad. Si su maestro le indica que en una prueba los puntajes se distribuyeron normalmente y que su puntaje se encuentra a 3 desviaciones estándar por encima del promedio, indudablemente usted tendría uno de los mejores puntajes.

Correlación: Nos permite observar posibles relaciones entre 2 o más variables. Si una correlación es de -4.13 tiene que haber un error de cálculo, ya que éstas deben encontrarse entre $+1, -1$.

Correlación para Escalas Nominales: No hay fluctuación de positivo a negativo. El resultado solo puede señalar la intensidad de la relación, pero no la dirección.

- **Comparaciones para tablas 2x2:** Esto significa de 2 variables con 2 alternativas cada una.
- 1. Phi ϕ : fluctúa entre 0 y 1. Según el coeficiente se acerca a 0, más baja será la relación; según se acerca a 1, más intensa será la relación.
- 2. Lambda λ : fluctúa entre 0 y 1; a más cerca del 1 mayor intensidad de la relación, y a más se acerca a 0 menor intensidad de la relación.

Ejemplo:

En UCA se tomó una muestra de profesores de los programas de psicología clínica, se clasificaron en términos del programa al cual pertenecen, y se obtuvo de estos su actitud en relación a la frase “que se vaya la Marina en Vieques”. Se obtuvo los siguientes datos

Programa al que pertenece

Actitud	Ph.D.	Psy.D.	Total
Favorable	30	39	69
Desfavorable	20	21	41
Total	50	60	110

Resultado: $\phi = .0326$ Lambda = 0

Interpretación de ϕ : no se encuentra una relación entre pertenecer a un programa en particular y presentar una actitud a favor o en contra de la Marina.

- **Coeficiente para tablas mayores de 2x2:** Significa que por lo menos una de las variables tiene 2 o más alternativas. Aquí se utiliza el coeficiente V de Cramer. También se puede utilizar el coeficiente de Lambda.

Ejemplo:

En la situación anterior se arregló considerando la afiliación política de la facultad. La distribución quedó de esta manera:

Afiliación Política

Actitud	PIP	PPD	PNP	Ninguno	Total
Favorable	28	25	9	7	69
Desfavorable	2	5	16	18	41
Total	30	30	25	25	110

Resultado: $V = .59$

Interpretación: el resultado nos sugiere una relación moderada alta entre la afiliación política de la facultad y la actitud hacia la presencia de la marina en Vieques. Puede notarse que los facultativos del PIP y del PPD muestran una actitud más favorable a la frase mientras que los del PNP y los afiliados a ningún partido tienden a desfavorecer este “slogan”.

Lambda = .439

Interpretación: el índice obtenido nos indica que el por ciento de reducción en error en la predicción de la actitud, conociendo la afiliación política de la facultad es de 43.9

Correlación para Escalas Ordinales: Es posible observar relaciones directas o inversas, siempre y cuando la relación sea lineal (y no curvilínea).

- **Spearman (Rho):** Es el coeficiente más utilizado para estas. Correlaciona rangos o posiciones, no la magnitud de las variables.
 1. **Correlación Positiva:** Entre dos pruebas significa que cuando un puntaje aumenta el otro también aumenta. Significa que hay una relación directa. A mayor ocurrencia de X evento, mayor probabilidad de que ocurra Y evento.
 2. **Correlación Negativa:** A mayor ocurrencia de X evento menor ocurrencia de Y evento y se interpretará como inversa. Ej. Un investigador encontró una alta correlación negativa entre fumar cigarrillos y la capacidad para apagar velas de su cumpleaños esto significa que las personas que más cigarrillos fuman pueden apagar menos velas.
- **Gamma:** Fluctúa entre 0 y 1 si la relación es directa y entre 0 y -1 si la relación es inversa.

Ejemplo:

Un investigador estaba interesado en determinar si había relación entre el alcohol y satisfacción marital. Utilizó 2 instrumentos de escala ordinal para operacionalizar las variables y obtuvo lo que aparece a continuación, donde a más alto índice mayor consumo de alcohol y mayor insatisfacción marital.

Sujeto	Nivel de Consumo de Alcohol	Nivel de Satisfacción Marital
1	7	6
2	4	3
3	8	7
4	5	4
5	9	8
6	10	9
7	6	6
8	3	2
9	2	1
10	4	3
11	1	2
12	5	4

Resultado: $r_s = .985$

Interpretación: el índice obtenido sugiere una relación directa sumamente fuerte entre el nivel de consumo de alcohol y la puntuación en la escala de satisfacción marital. Puede encontrarse una tendencia de a mayor consumo de alcohol mayor insatisfacción marital.

Gamma = .96

Sugiere una intensa relación directa entre el consumir alcohol y satisfacción marital (se interpreta de forma similar).

Correlación para Escalas de Intervalos/Razón: El índice de correlación que se aplica se denomina correlación producto momento de Pearson (r). Esta solo puede utilizarse cuando la relación es lineal. Directa 0 y 1; Inversa 0 y -1.

Ejemplo:

Un psicólogo industrial tomó una muestra de empleados y determinó el número de días que se ausentaron durante 6 meses, así como el nivel de involucración en el trabajo. Obtuvo lo siguiente:

Sujeto	A	B	C	D	E	F	G	H	I	J	K	L
Involucración	40	60	85	55	70	65	85	90	45	50	80	50
Ausencias	10	8	3	9	7	6	2	1	9	7	2	6

Resultado: $r = -.91$

Interpretación: la correlación obtenida sugiere que entre las ausencias y la involucración al trabajo hay una relación inversa fuerte, esto es, a mayor involucración, menor ausencias.

Rbp: Correlación biserial puntual. Una variable Nominal y una variable de Intervalo/Razón.

Análisis de Reactivos

Índice de dificultad o facilidad

Dificultad: Seleccionar ítems con P de 25% a 75%

Categorías para P:	Categorías para Delta (Δ)
1.0 - .86 Demasiado Fácil	6 – 11 Fácil
.85 - .71 Fácil	12 – 14 Moderado
.70 - .31 Promedio	15 – 19 Difícil
.30 - .16 Difícil	
.15 - .00 Demasiado Difícil	

Índice de Discriminación

Discriminación: Seleccionar ítems con D de .30 en adelante

Categorías para D:

- $\geq .40$ = ítem muy bueno
- .30 - .39 = ítem razonablemente bueno, sujeto a revisión
- .20 - .29 = ítem marginal, necesita revisión
- $\leq .19$ = ítem pobre, se debe rechazar o revisar

Categorías para r_{bis} (índice de correlación biserial):

- 1.0 - .70 Muy alto
- .70 - .30 Alto (discrimina de forma adecuada)
- .29 - .00 Muy bajo

Coefficiente Eta: Llamado algunas veces *razón de correlación*, se utiliza cuando una relación bivalente es principalmente curvilínea. Aplica tanto a escalas cualitativas como cuantitativas.

Regresión: Es el análisis de las relaciones entre variables.

- **Regresión Simple:** Implica una variable independiente (X) y una variable dependiente (Y).
- **Regresión Múltiple:** Supóngase que el director sospecha que la predicción del promedio de calificaciones mejorará si se usa también como pronosticador otra puntuación de prueba, digamos, una puntuación en una prueba de habilidades motoras finas. El uso de más de una puntuación para predecir Y requiere el uso de una ecuación de regresión múltiple. Es una extensión de la regresión simple a las situaciones con dos o más variables de predicción.

Un **intervalo de confianza** es un rango de valores que probablemente contiene el valor poblacional. El intervalo de confianza del 95% es aquel en el que la probabilidad de que dicho intervalo contenga el valor poblacional es de 0.95. Los intervalos más utilizados en la práctica son el de 95 y 99 por ciento.

Los **límites de confiabilidad** son los valores que establecen las fronteras del intervalo de confianza.

ESTADÍSTICA INFERENCIAL

Estimación de Parámetros: Se realiza sobre las muestras, para estimar el nivel de una o más características poblacionales; por ejemplo, el promedio o porcentaje de una población. Las encuestas de opinión pública, los escrutinios y gran parte de la investigación de mercado.

- **Estimación de punto:** Por ejemplo: Supongamos que la administradora de una universidad quiere conocer el IQ promedio de los profesores que laboran en ese centro educativo. La administradora extrae una muestra aleatoria y el promedio es igual a 135. La estimación sería, entonces, de 135; este número es una estimación puntual, ya que se utiliza solamente un valor de estimación.
- **Estimación de intervalo:** Si le preguntamos a la anterior lo que piensa respecto a que la media de la población es exactamente igual a 135, seguramente respondería que no lo creería. Pero entonces, ¿Qué tan cerca está 135 de la media de la población?. La forma usual de responder esta pregunta es dar un rango de valores que con certeza incluyan a la media de la población. Esta es una estimación de intervalo. Por ejemplo, la administradora podría tener cierta confianza en que la media poblacional se encuentre en el rango 130 – 140.

Hipótesis Científica: Aseveración que supuestamente explica o soluciona el problema. Casi nunca se pueden probar directamente. Es necesario para definir variables operacionales. No pueden probar causa y efecto.

Hipótesis Estadística: Se operacionaliza la hipótesis científica y se definen las variables. Son enunciados respecto a uno o más parámetros. Es una expresión de la hipótesis científica en términos de afirmaciones respecto a parámetros desconocidos.

- **Hipótesis Nula:** La usamos cuando comparamos grupos y estos no difieren. Establece que son los factores aleatorios los que explican las diferencias que se observan entre los grupos que se comparan y no la variable independiente.
- **Hipótesis Alternativa:** Nos dice si existen diferencias entre ambos grupos. Sostiene que las diferencias son significativas. Que las discrepancias no están explicadas por factores aleatorios sino por la variable independiente que introduce o examina el investigador. Es la que mejor representa la hipótesis científica.
 1. *Unidireccional:* Sustenta que existen diferencias significativas en una dirección dada. Se utiliza cuando se conoce el problema.
 2. *Bilateral:* La misma afirma la existencia de diferencias significativas, sin especificar la dirección en que se esperan dichas diferencias.

Error Alfa o Tipo 1: Rechaza la Hipótesis Nula cuando es cierta y debió aceptarse. Se usa .001 para que tenga menos probabilidad de rechazar.

Error Beta o Tipo 2: Se Acepta la Hipótesis Nula cuando es falsa.

Probabilidad: La frecuencia en que se manifiesta cada posibilidad.

Prueba T de Wilcoxon (Tw): Se usa para diseños intragrupos. Considera la magnitud del cambio observando antes y después de una prueba en la escala ordinal. La muestra debe estar compuesta entre 5 y 50 participantes. Se analiza cambios Después – Antes, se asignan los signos (+ ó -) y rango.

Prueba Mann-Whitney (U): Comparación de 2 muestras independientes en escala ordinal. Muestras entre 3 y 20 casos. Se usa entre grupos. Diferencias entre los rangos de dos grupos. Si la muestra es mayor de 20 se hace un ajuste para usar Z.

Prueba de t: Se utiliza con escalas de intervalo y de razón. Nos dice si el comportamiento promedio de dos grupos distintos son significativamente diferentes y cuán grandes son las diferencias entre los promedios. La prueba *t* pretende eliminar la variación interna, ya sea en grupos independientes (aleatorizados) o pareados (dependientes). Esta prueba busca diferencias en el comportamiento de los grupos. También nos ayuda a rechazar o aceptar la hipótesis. Menos de 30 participantes.

Curva Normal Z: Comparación del promedio o la proporción de una muestra con el universo. La muestra debe estar compuesta de 30 participantes o más. Se debe conocer la desviación estándar del universo. La Z obtenida se compara con un valor crítico dado el nivel de significación y su hipótesis estadística. Si la obtenida es mayor que el valor crítico hay diferencias significativas.

Chi Cuadrada (X²): El coeficiente de correlación X² es una medida de discrepancia entre lo que se observa y lo que se espera. Establece cuán significativas son las

diferencias entre los miembros de un grupo y otro. Es una prueba no paramétrica y equivalente al coeficiente Phi. No obstante, se aplica con variables multicatóricas de dos o más niveles. Se utiliza un modelo matemático con lo que se obtiene el valor crítico de X^2 para su interpretación en función a la probabilidad Alfa establecida previamente.

Anova: Varianza entre grupos deben ser iguales u homogéneos. Variable Dependiente tiene que ser continua y normalmente distribuida. Mientras que la Variable Dependiente tiene que ser categórica o nominal. Se puede usar con dos o más Variables Independientes esto se llama *Anova Factorial* si se cruzan hay interacción, si no se cruzan no hay interacción.

Por ejemplo: Ropa más blanca con ACE o Ariel o con agua fría y caliente. Si cada una ponen la ropa más blanca o intercalado una con otra la pone más blanca.

ACE o Ariel = independientemente el detergente

Agua Caliente o Fría = independientemente el agua

Agua fría Ace-Ariel

Agua caliente Ace-Ariel.

Problema de Investigación: Consiste en la idea de investigación formulada en términos concretos y explícitos, susceptible de investigación con procedimientos científicos. El problema debe implicar la posibilidad de realizar una prueba empírica, de poder observarse y medirse. Consiste de los siguientes elementos: Objetivos de la Investigación = Qué voy a estudiar y por qué?; Preguntas de Investigación = Cada área de estudio estructurándola parte x parte y cada pregunta contesta un objetivo o una hipótesis reafirmada.; Justificación = Por qué vale la pena llevar a cabo esta investigación.

Diseños Pre-Experimentales: La investigación en la cual se utiliza un diseño pre-experimental, a pesar de que se siguen parámetros rigurosos en la aplicación de tratamientos, no cumplen con toda la rigurosidad del control de las variables, en específico, la selección aleatorizada de participantes. Sin embargo, estos estudios son útiles para establecer la base para llevar a cabo otras investigaciones. Estos diseños se utilizan mucho, por ejemplo, en el área de educación.

- **Estudios de un Caso con una Sola Medición:** En este no hay grupo de comparación y no hay control. Tiene poco valor científico. $X \quad O$
- **Diseño pre-prueba/pos-prueba de un solo grupo:** Se mide la condición experimental antes del tratamiento y después del tratamiento, para comparar los resultados. $O_1 \quad X \quad O_2$
- **Diseño de un grupo estático:** Se administra un tratamiento (X) a un grupo (O_1) y se compara con otro grupo (O_2) que no ha recibido el tratamiento para establecer su posible efecto. La línea entrecortada significa que el grupo no es equivalente o no son iguales porque no fueron aleatorizados.

Diseños Experimentales: La investigación experimental se desarrolla en ambientes rigurosamente controlados, principalmente en laboratorios. Estos diseños gozan del mayor prestigio por la exactitud en las observaciones y medidas que se toman en su desarrollo. El investigador manipula la variable independiente, considerada como la causa y observa el cambio en la variable dependiente, considerada el efecto. Por los

general, estos diseños utilizan ecuaciones basadas en símbolos para ilustrar los procedimientos a ser aplicado al Grupo Contro y Grupo Experimental.

Diseño ExPost-Facto: En este tipo de investigación, los datos se registran después de que el hecho ha ocurrido. El investigador no manipula la variable independiente directamente.

$$\begin{array}{c} X \quad O_1 \text{ Grupo Experimental} \\ R \text{ -----} \\ \quad O_2 \text{ Grupo de Comparación} \end{array}$$

Diseño Factorial: Se considera el modo en cómo se manipulan o se controlan las variables. Estos controles definen si el diseño es de una categoría experimental o sencillamente de naturaleza descriptiva. El diseño factorial facilita el estudio simultáneo de dos o más variables y niveles o valores de la Variable Independiente. Cada nivel de la VI constituye una condición experimental a ser considerada. Mediante este diseño se registran la interacciones entre variables y se analizan los efectos de las variables extrañas en vez de controlarlas según sea el caso. La ecuación dependerá del número de variables independientes y niveles o valores de cada variable a ser estudiadas.

Ejemplo:

VI: Cantidad de Esfuerzo Físico		
	Nivel 1: Mucho	Nivel 2: Poco
Nivel 1: Mucho	Mucho Esfuerzo Mucho Material	Poco Esfuerzo Mucho Material
Nivel 2: Poco	Mucho Esfuerzo Poco Material	Poco Esfuerzo Poco Material

VI: Cantidad de Material aprendido a través de la Lectura

VD: Interacciones o efectos

Diseños Cuasi-Experimentales: El investigador no tiene total control de sus variables. La falta de control amenaza la validez interna y externa del experimento. Tales variables son las llamadas Variables Extrañas o interventoras. No obstante, es una alternativa para investigar ciertas variables. Se utiliza frecuentemente en investigaciones en el área de educación.